



DEUTSCHES
PATENTAMT

12 Übersetzung der
europäischen Patentschrift

51 Int. Cl.⁶:
G 06 F 3/00

87 EP 0 554 492 B1

10 DE 692 04 045 T 2

- | | | |
|----|---|--------------|
| 21 | Deutsches Aktenzeichen: | 692 04 045.5 |
| 86 | Europäisches Aktenzeichen: | 92 102 049.1 |
| 86 | Europäischer Anmeldetag: | 7. 2. 92 |
| 87 | Erstveröffentlichung durch das EPA: | 11. 8. 93 |
| 87 | Veröffentlichungstag
der Patenterteilung beim EPA: | 9. 8. 95 |
| 47 | Veröffentlichungstag im Patentblatt: | 18. 4. 96 |

73 Patentinhaber:

International Business Machines Corp., Armonk,
N.Y., US

74 Vertreter:

Teufel, F., Dipl.-Phys., Pat.-Ass., 71155 Altdorf

84 Benannte Vertragsstaaten:

DE, FR, GB

72 Erfinder:

Korth, Hans-E., Dipl.-Phys., W-7000 Stuttgart, DE

54 Verfahren und Vorrichtung zum optischen Eingang von Befehlen oder Daten.

Anmerkung: Innerhalb von neun Monaten nach der Bekanntmachung des Hinweises auf die Erteilung des europäischen Patents kann jedermann beim Europäischen Patentamt gegen das erteilte europäische Patent Einspruch einlegen. Der Einspruch ist schriftlich einzureichen und zu begründen. Er gilt erst als eingelegt, wenn die Einspruchsgebühr entrichtet worden ist (Art. 99 (1) Europäisches Patentübereinkommen).

Die Übersetzung ist gemäß Artikel II § 3 Abs. 1 IntPatÜG 1991 vom Patentinhaber eingereicht worden. Sie wurde vom Deutschen Patentamt inhaltlich nicht geprüft.

DE 692 04 045 T 2

DE 692 04 045 T 2

B E S C H R E I B U N G

Verfahren und Vorrichtung zur optischen Eingabe von Befehlen oder Daten

Die vorliegende Erfindung liegt im Bereich von Schnittstellen zur direkten Eingabe von Daten oder Befehlen durch den Benutzer. Genauer gesagt, betrifft die Erfindung Geräte, die manuell bedient werden und auf Fingerdruck oder andere Handbewegungen reagieren.

Um Daten beispielsweise in einen Rechner einzugeben, ist die Tastatur das am meisten verwendete Gerät. Sie ermöglicht eine schnelle, eindeutige Eingabe von digitalen Daten. Alternative Eingabegeräte können für besondere, begrenzte Anwendungen verwendet werden: Spracherkennung ist inhärent schnell, erfordert aber einen sehr großen Verarbeitungsaufwand. Die Umwandlung von Sprache in Text schließt zwangsläufig Mehrdeutigkeit ein. Eine automatisierte Analyse der Handschrift ist auch schwierig: Die Eingabe von leicht lesbaren einzelnen Zeichen ist langsam und umständlich. Das Lesen individueller Handschriften kann eine schwierige und zeitraubende Aufgabe sein.

Zeigegeräte (Maus, Stift) haben breite Anwendung in Verbindung mit Grafikumgebungen in Fenstertechnik gefunden, um die Verarbeitung zu steuern und um Binärdaten einzugeben. Bei dieser Anordnung erfordert eine typische Datenverarbeitungssitzung ein häufiges Wechseln zwischen Zeiger und Tastatur.

In CAD-Anwendungen wird typischerweise eine Maus zur Editierung von Grafiken benutzt. Andere Geräte, z.B. Joysticks oder Rollkugeln, können zum System hinzugefügt werden, um eine komfortable Eingabe von Anzeigebefehlen zu ermöglichen, wie beispielsweise die Rotation eines virtuellen 3D-Objekts zur Erzeugung einer Zeichnung aus einer anderen Perspektive.

Ein bestimmtes optisches Eingabegerät zur Verwendung in einer Grafikstation ist in IBM TDB, Jahrgang 32, Nr. 3B, August 1989, Seite 92 beschrieben. Dieses Gerät besteht aus einem Hand-Würfel, dessen Seiten eine bestimmte Farbe haben, um Anhaltspunkte für die Orientierung auf der Oberfläche zu geben. Die aktuelle Position und Ausrichtung des Würfels wird von einem TV-Sensor erkannt. Das aufgezeichnete Bild des Würfels wird digital verarbeitet, und Befehle werden entnommen, um die grafische Darstellung eines virtuellen 3D-Objekts entsprechend der Ausrichtung des Würfels zu ändern. Dieser Technik sind bleibende Probleme inhärent, da die Orientierungshilfen aufgrund dessen, daß der Würfel in der Hand gehalten wird, teilweise verdeckt sind.

Mit dem anhaltenden Trend zur Miniaturisierung sind moderne Rechner durch verringerte Abmessungen gekennzeichnet. Während die Größe der Elektronikbauelemente abnimmt, können die Abmessungen der Benutzerschnittstelle, z.B. der Tastatur oder der Maus, nicht zu sehr reduziert werden, ohne Einbußen bei der Funktionalität hinzunehmen.

JP-A-59132079 beschreibt ein manuell bedienbares Eingabegerät, um die Eingabe durch Zeigen möglich zu machen, ohne daß der Bediener Hardware in die Hand nehmen muß, indem die Form einer manuellen Bedienung als ein Bild eingegeben und dies unterschieden und verarbeitet wird. Die Form und die Position der Hand eines Menschen werden von einem Bildeingabeteil als zweidimensionale Bildinformationen eingegeben. Diese zweidimensionalen Bildinformationen werden an einen die Form und die Position trennenden Teil übertragen. Die Form der Hand und der Finger, ihre relativen Positionen auf einem Bildschirm und ihre absoluten Positionen werden getrennt, und Informationen über die Form und die Position erhält man durch Vergleich mit dem Hintergrund. Diese Informationen über die Form und die Position werden an einen Merkmale entnehmenden Teil übertragen, und individuelle Kennzeichen, wie beispielsweise ein individueller Unterschied der Größe der Hand und der Finger, werden eliminiert, und die Merkmale von Bedienungsformen einer universellen Hand und von

universellen Fingern werden entnommen. Die auf diese Weise entnommenen Merkmalsinformationen von Bedienungsformen der Hand und der Finger werden an einen Anzeigesignal-Ausgangsteil übertragen. Die vorläufig festgestellten Merkmale von Bedienungsformen der Hand und der Finger und Anzeigeinformationen werden mit dem Inhalt einer Umsetzungstabelle verglichen und in Anzeigeinformationen an ein Verarbeitungsgerät umgewandelt und ausgegeben.

JP-A-2132510 beschreibt ein Eingabegerät, das es direkt ermöglicht, den Willen eines Menschen widerzuspiegeln, indem die Rechnereingabe entsprechend der Position von beispielsweise einer Hand erreicht wird, die von einer Videokamera aufgenommen wird. Eine Hand wird zunächst geschlossen auf eine Anzeigeplatte gelegt, dann wird das Bild der Hand mit ihr überlagert angezeigt. Dann wird ein Cursor auf dem Bild entsprechend der Position der Mittenkoordinate des ersten der Bilder angezeigt. Ein Befehl wird ausgewählt, wenn die Hand auf der Platte bewegt wird, und der auf das Bild gesetzte Cursor wird auf ein Symbol bewegt. Wenn ein Finger der Hand beispielsweise geöffnet wird, wird die Spitze des Fingers erkannt und gedrückt. Dann wird ein von dem Symbol ausgewählter Befehl eingegeben. Auf diese Weise ist es möglich, die direkte Rechnereingabe ohne den Einsatz von speziellen Geräten vorzunehmen und den Willen eines Menschen widerzuspiegeln.

Die vorliegende Erfindung ist dazu gedacht, die Nachteile zu beseitigen, die den dem Stand der Technik entsprechenden Benutzereingabegeräten inhärent sind.

Es ist daher eine Aufgabe der vorliegenden Erfindung, eine neue und weiterentwickelte Schnittstelle zur Eingabe von Daten oder Befehlen in Daten- oder Befehlsverarbeitungssysteme bereitzustellen, die Probleme vermeidet, die mechanischen Bauelementen anhaften, und es ermöglicht, verschiedene Arten von manuell bedienten Eingabegeräten zu emulieren.

Dies wird von der vorliegenden Erfindung, wie beansprucht, er-

reicht, die eine optische Benutzerschnittstelle beschreibt, bei der es sich im allgemeinen um eine Anordnung handelt, die ein Bilderfassungssystem einsetzt, um die Hand- und Fingerbewegungen und die Gesten eines Benutzers zu überwachen. Diese Aktivitäten werden als Operationen auf einer physisch nicht vorhandenen Rechnertastatur oder einem anderen Eingabegerät interpretiert. Das Bilderfassungssystem umfaßt einen TV-Sensor und Mittel zur Digitalisierung und Verarbeitung von Signalen von dem TV-Sensor. Hand, Finger und Fingerspitzen des Benutzers werden von dem TV-Sensor überwacht, und aktive Bewegungen werden von den Mitteln für die Signalverarbeitung festgestellt und analysiert und als entsprechende Eingabedaten oder Befehle interpretiert.

Die erforderliche Hardware für die erfindungsgemäße optische Schnittstelle ist lediglich ein (z.B. ein CCD-Wandler-) Videosensor mit einer "Bildabtast"-Schaltung. Eine 100x100-Pixelmatrix sollte das Mindeste für die erforderliche Bildverarbeitung sein. Eine anamorphotische Fotolinse oder ein Sensorchip mit gedehnten Pixeln kann verwendet werden, um den Sensor auf das Seitenverhältnis der Tastatur abzustimmen. Autofokus oder Fixfokus vereinfachen die Einstellprozedur. Standard-Sensoren für Camcorder oder elektrophotografische Sensoren erfüllen diese Voraussetzungen problemlos.

Eine Schwarz/Weiß-Abbildung ist ausreichend, wenn sich die Arbeitsoberfläche durch eine Graustufe unterscheiden läßt, die sich wesentlich von der Tönung der menschlichen Haut abhebt. Wenn sich keine Vermutungen über den Hintergrund (z.B. die Schreibtischoberfläche) anstellen lassen, ist eine Vorstufen-Bildverarbeitung erforderlich, um die Konturen der Hände des Bedieners festzustellen. Mit dem a-priori-Wissen, daß das Hintergrundbild konstant bleibt (Rauschen und Schattierung ausgenommen), und dem Wissen über die Hand als eine zusammenhängende Einheit, hält sich das notwendige Maß an zeitkritischer Verarbeitung in Grenzen.

Es sei erwähnt, daß die Kosten für die Kamera und für die Hard-

ware zur Bildabtastung aufgrund der geringeren Anzahl von Teilen letztendlich niedriger sein können, als die Kosten für eine gute, allein funktionsfähige Tastatur mit Ansteuerschaltung. Eine weitere Kostensenkung ergibt sich aus der Tatsache, daß einige defekte Pixel (beispielsweise auf dem Ladungsspeicherchip) toleriert werden können.

Eine TV-Kamera mit 50 (Halb-)Bildern pro Sekunde ist schnell genug, um Tippgeschwindigkeiten von ca. 10 Tastenanschlägen pro Sekunde zu überwachen. Während allgemeine Bildverarbeitungsaufgaben eine hohe Rechenleistung erfordern, lassen sich Tastenanschläge recht wirtschaftlich feststellen, da dies viel a-priori-Wissen einschließt.

Wird die Kamera beispielsweise von einem Monitor oder einem Rechnergehäuse in Richtung des Bedieners gelenkt (Figur 1), ist der untere Teil des Bildes vom Arbeitstisch ausgefüllt. Objekte, die sich über den Tisch bewegen, können ganz einfach anhand ihrer Größe, Form und Bewegung als Hände erkannt werden. Wenn eine Hand erkannt worden ist, ist ihre Positionsabweichung in aufeinanderfolgenden Bildfeldern ziemlich gering. Daher muß nur eine relativ geringe Anzahl von Pixeln ausgewertet werden, um die Bewegung zu verfolgen und um die Vollständigkeit der Bildszene zu prüfen.

Unter der Annahme der Form einer Hand über einer imaginären Tastatur erlaubt ein einfaches Modell, die Positionen der Fingerspitzen den (bis zu fünf) hervorstehenden Stellen an der Unterseite der Handkontur zuzuordnen. Die Bewegungen der Finger in bezug auf die Hand sind ebenfalls klein. Sie können durch die Auswertung von kleinen Bildbereichen effizient überwacht werden. Die Stetigkeit der Bewegungen erlaubt es, die Positionen der Suchbereiche zu extrapolieren.

Es ist klar, daß die Bewegungen und die Form der Hände zusätzlich zu den Bewegungen der Fingerspitzen dazu verwendet werden können, Gesten mit Zusatzinformationen an die optische Schnitt-

stelle zu senden (z.B. Beginn/Ende der Sitzung, Tastatur-Neubelegung, Kameramodus oder andere Befehle). Somit realisiert die Erfindung ein weiterentwickeltes berührungsloses Daten-/Befehls-eingabegerät, das in der Lage ist, manuell bediente Mittel in Datenverarbeitungssystemen oder elektronischen Verbindungen im allgemeinen zu ersetzen.

Die optische Benutzerschnittstelle kann für eine Vielzahl von Aufgaben verwendet werden. Sie können im Nu gewechselt werden, und zwischen ihnen und der aktuellen Anwendung kann umgeschaltet werden.

Alphanumerische Tastaturen: Ein beachtlicher Vorteil bietet eine virtuelle Tastatur für Rechner, die kleiner sind als ein Notebook. Eine Tastatur in voller Größe ist mit dem Kompaktgerät erhältlich.

Tastaturen für große Zeichensätze: Die visuelle Darstellung der aktuellen Tastenfunktion erlaubt die Handhabung von Tastaturen mit mehreren Umschalttasten und die Anzeige von Untergruppen für große Zeichensätze (z.B. Kanji).

Rechenfelder: Die Anordnung kann optimiert werden, beispielsweise für die Eingabegeschwindigkeit, Einfachheit oder eine entsprechende Reihe von Funktionstasten.

Arbiträre analoge und digitale Bedienfelder: Verschiedene Bedienfelder, z.B. Rollkugel, Joystick, können für die aktuelle Aufgabe emuliert werden.

Tastaturen für Musikinstrumente: Ausgestattet mit einer virtuellen Klaviertastatur wird aus einem PC ein vollständiges System zur Musikproduktion. Externe Tastaturen, Synthesizer/Folgesteuerungseinheiten werden der Vergangenheit angehören. Eine anschlagsdynamische Empfindlichkeit ist dem Sensor-Algorithmus inhärent. Mit Kamera und Klaviertastatur hat ein PC eine "Multi-media"-Schnittstelle, welche die Eingabe, Ausgabe, Verarbeitung

und Speicherung von Bildern und Melodien ermöglicht.

Identifizierung der persönlichen Unterschrift: Eine sichere Identifikation läßt sich aus der Unterschriftsdynamik erhalten.

Alternative Kameraeingabe: Die Bilderfassungsfunktion des Systems kann zusätzlich für andere Zwecke verwendet werden, z.B. das Abtasten von Dokumenten, FAX, Bildtelefon.

Für ein besseres Verständnis der vorliegenden Erfindung, zusammen mit weiteren Aufgaben und Vorteilen, ist als eine bevorzugte Ausführungsform der Erfindung im folgenden die Verwendung der optischen Schnittstelle als virtuelle Tastatur mit Bezug auf die Begleitzeichnungen beschrieben, wobei:

Fig. 1 die Anordnung des Bilderfassungssystems zeigt, um die Hand- und Fingerbewegungen eines Benutzers zu erfassen.

Fig. 2 eine schematische Abbildung des Kamerabildes zusammen mit einer verbesserten Bereichsaufteilung einer überwachten Hand ist.

Fig. 3 die Schritte der Verarbeitung des Bildes einer Hand aufzeigt.

Eine Tastatur hat die Aufgabe, Informationen über die vorhandenen Tasten und ihre Position zu geben, das Drücken von Tasten durch den Bediener zu erfassen und eine Rückmeldung zu erzeugen, wenn eine Taste berührt worden ist.

Die Tastatur braucht nicht physisch vorhanden zu sein. In Fig. 1 dient eine virtuelle Tastatur 3 als Daten- oder Befehlseingabegerät. Von der Rechnerseite der Schnittstelle muß es lediglich eine Möglichkeit geben, die Bewegungen der Fingerspitzen des Tastaturbedieners zu erkennen. Herkömmlicherweise erfolgt dies mit elektrischen Kontakten unter den einzelnen Tasten. Andere

Ansätze jedoch funktionieren ebenso: Wenn die Hand- und Fingerbewegungen auf einer virtuellen Tastatur (z.B. einer Tischtastatur) von einem optischen System überwacht werden, kann die geeignete Bildverarbeitung die Absicht des Benutzers erkennen.

Fig. 2 zeigt den von der Kamera 2 erfaßten Bereich. In dem überwachten Bereich 4 werden die Hände des Benutzers erkannt. Der vergrößerte Teil 5 von Fig. 2 stellt einen Bereich des Kamerabildes dar, der für die weitere Bildverarbeitung gewählt wird.

Ein dreistufiger Algorithmus minimiert die Ausführungszeit (Fig. 3). Der Algorithmus der ersten Stufe fragt Kontrastwerte auf einer Linie 6 ab, um das Vorhandensein einer Hand festzustellen. Die Routine der zweiten Stufe folgt der Handkontur 7. Der Algorithmus der dritten Stufe stellt die Positionen der Fingerspitzen (81, 82, 83, 84, 85) fest.

Informationen über die Tasten für eine virtuelle optische Tastatur 3 (VOK) können in einem Fenster auf der Monitoranzeige 1 angezeigt werden. Schematische Darstellungen der Bilder der Hand können über einer Tastaturschablone angezeigt werden. Die gleichzeitige Anzeige der Tastatur mit der Anwendung ist aus ergonomischen Gründen vorteilhaft, da man den Blick zwischen Bildschirm und Tastatur nicht zu wechseln braucht.

Das Tastaturfenster erfordert vielleicht ca. 10 % bis 20 % des Anzeigebereichs. Jedoch kann das Fenster nur während der Dateneingabe erscheinen, so daß andere, gleichzeitige Anwendungen nicht betroffen sind. Beim Prüfen der Eingabeposition des Cursors positioniert sich das Tastaturfenster auf dem Bildschirm gegebenenfalls gegenüber des Texteingabebereichs.

Eine Fingerspitze, die eine virtuelle Taste anschlägt, macht eine schnelle Bewegung nach unten (über einen Weg von ca. 10 mm), hält abrupt an, wenn die Schreibtischoberfläche getroffen ist und bewegt sich nach kurzer Entspannung wieder nach oben. Die Stelle, an der die anschlagende Fingerspitze anhält, kann

einer Taste zugeordnet werden. Mehrdeutigkeiten aufgrund von sich gleichzeitig bewegendem Fingern lassen sich anhand einer Auswertung der unterschiedlichen Bewegungscharakteristika (der anschlagende Finger bewegt sich schneller über einen langen Weg und hält abrupt an) sowie anhand eines Handmodells beseitigen, das unwahrscheinliche relative Fingerpositionen aufzeigt.

In der Praxis läßt sich die Position einer Fingerspitze für jedes Bildfeld (oder Halbbild) bestimmen, d.h. ca. alle 40 msec. Eine Folge von fünf Bildfeldern erlaubt die Überwachung der Bewegung der Fingerspitze über einen Zeitraum von 160 msec. Die unterschiedliche Position der Fingerspitze zwischen zwei aufeinanderfolgenden Bildfeldern ergibt dann vier Werte für die Geschwindigkeit des Fingers ($V_4 \dots V_1$). Um einen Anschlag festzustellen, muß eine Folge von Geschwindigkeitswerten gefunden werden, die der obigen Beschreibung entspricht. Um dieses Problem mit nur vier Geschwindigkeitsmessungen zu lösen, soll von einigen Annahmen ausgegangen werden:

- Vor dem Anschlag ist die Beschleunigung der Fingerspitze relativ konstant. Dies scheint vernünftig zu sein, da die Beschleunigung lediglich von der Muskelspannung abhängt. In diesem Fall kann die Geschwindigkeit (V_x) in der Nähe des Anschlagpunktes durch Extrapolation aus den beiden Geschwindigkeitswerten vor dem Anschlag berechnet werden. Der extrapolierte Wert kann dann mit einem Schwellenwert verglichen werden.
- Der Anschlag einer virtuellen Taste schließt ein, daß die Geschwindigkeitswerte vor dem Anschlag über einem Mindestwert liegen müssen.
- Beim Anschlag nimmt die Geschwindigkeit rapide ab. Daher enthält die Geschwindigkeitsmessung während des Anschlags (V_2) keine hilfreichen Informationen. Sie kann jedoch dazu verwendet werden, die Gültigkeit der Meßfolge zu prüfen (kein Rauschen).

- Nach dem Anschlag läßt die Geschwindigkeit der Fingerspitze rapide nach. Die Geschwindigkeitsabnahme ist durch die Deformation der elastischen Fingerspitze bedingt. Die Geschwindigkeit nach dem Anschlag muß in einem Bereich liegen, der von der Mindestverzögerung und dem Maximum der folgenden erneuten Beschleunigung bestimmt wird.

Mit diesen Annahmen kann eine Reihe von vier einfachen Bedingungen festgeschrieben werden, die für einen Anschlag erfüllt werden müssen:

1. $V_x > V_{min}$ (V_{min} = Mindestanschlagsgeschwindigkeit)
2. $V_4 > V_p$ (V_p = Mindestgeschwindigkeit vor dem Anschlag)
3. $V_2 > V_1$ keine Schwingung
4. $V_1 < V_t$ (V_t = Höchstgeschwindigkeit nach dem Anschlag)

Wenn ein Tastenanschlag festgestellt worden ist, muß die entsprechende Taste gefunden werden. Dies ist keine triviale Aufgabe, da sich der Finger bewegt. Wenn der Finger visuell geführt wird, erfolgt die Tastenauswahl bevor die Anschlagsbewegung beginnt, d.h. ca. 100 msec bevor der Anschlag festgestellt wird. Für ein visuell gesteuertes, schnelles Tippen wird der Impuls "Suchen und Anschlagen" des Bedieners aus der sichtbaren Entfernung zwischen dem Finger und der anzuschlagenden Taste erzeugt. In diesem Fall ist es die virtuelle Tastaturschablone, auf welcher der aktuelle Anschlagspunkt ausgewählt werden sollte. In der Praxis muß ein stetiger Übergang zwischen diesen Extremen realisiert werden. Die Kenntnis des Weges der Fingerspitze ermöglicht es, das Kriterium der Geschwindigkeit vor dem Anschlag und der Beschleunigung für den Algorithmus zur Feststellung eines Anschlags zu verfeinern.

Eine selbstzentrierende Routine ist erforderlich, um die taktile Rückmeldung hinsichtlich der Tastenmitte zu übernehmen, die eine physische Tastatur bietet: Der Bediener möchte die Mitte der Taste treffen. Dies erlaubt einem Algorithmus, das Fingerbild entsprechend neu zu positionieren. Um eine unerwünschte Abwei-

chung des Handbildes zu vermeiden, muß eine Art "Steifheit" realisiert werden.

Um die Fingerspitzendynamik zu überwachen, werden Informationen über die protokollierten vertikalen Bewegungen einer jeden Fingerspitze gepuffert. Eine unterbrechungslose Folge von fünf Fingerspitzen-Messungen ist notwendig, damit ein Finger für den Anschlagsfeststellungs-Algorithmus berücksichtigt werden kann. Während der Überwachung kann sich jedoch die Anzahl der sicher festgestellten Fingerspitzen ändern (z.B. aufgrund von schlechter Beleuchtung, Fingern außerhalb des Videobild-Bereichs oder sich überlagernden Fingern). In diesem Fall kann sich die Fingerzahlzuordnung ändern, und das Fingerprotokoll ist verfälscht. Zur Vermeidung einer Totzeit von fünf Bildfeldern für einen neuen Aufbau der Protokollpuffer ist ein Neuuzuordnungs-Algorithmus erforderlich: Wenn die Fingeranzahl erhöht wird, müssen die neuen Finger erkannt werden. Zeiger müssen erzeugt werden, die den korrekten Zugriff auf die Protokollpuffer ermöglichen. Im Falle einer abnehmenden Fingeranzahl müssen die Zeiger für die verbleibenden Finger aktualisiert werden. Mit den Routinen der virtuellen Tastatur werden alle Finger (außer den neu hinzugekommenen) laufend überwacht. Selbst im Falle von Schwankungen, die durch Rauschen bei der Fingeranzahl bedingt sind, können die Finger für die Anschlagssignatur überwacht werden.

Eine Reihe von einfachen intuitiven "Gesten" kann definiert werden, um die Bedienung einer virtuellen Tastatur zu steuern, z.B.:

- Beginn/Ende: Der Bildverarbeitungs-Algorithmus prüft, ob eine Hand vorhanden ist, wobei er die "Handgelenkslinie" abfragt. Dies erfolgt beispielsweise alle 320 msec oder wenn ein Erkennungsfehler bei der Handkontur auftritt. Wird eine Hand erkannt, wird das Fenster mit der Tastaturschablone auf dem Bildschirm geöffnet. Wenn die Hände zurückgenommen werden, wird das Fenster geschlossen.

- Grundstellung bei der Tastatureingabe: Die Platzierung einer Hand auf dem Schreibtisch (d.h. der Anschlag mit fünf Fingern gleichzeitig) kann als Befehl interpretiert werden, die virtuelle Tastatur so zu platzieren, daß sich beispielsweise die linke Hand über den Tasten des A-S-D-F-Raums befindet.
- Rückgängig machen: Eine schnelle, nach oben gerichtete Fingerbewegung nach einem Tastenanschlag ist ein einfaches Signal zur "Rückgängigmachung" des Anschlags. Höchst intuitiv ist die Überwachung der Daumen. Unbewußt werden die Daumen durch die überraschte Reaktion auf einen Tippfehler nach oben gerichtet. Die Geste "rückgängig machen" vereinfacht die Tastaturbetätigung: Die Unterbrechung der momentanen Arbeit, um die Rückschritt- oder die Löschtaste zu finden, kann vermieden werden.
- Moduswechsel: Verschiedene Gesten mit einer oder beiden Händen können als Wechsel zwischen verschiedenen Tastaturmodi verstanden werden (z.B. Schreibmaschine, Maus, Taschenrechner, Finger Stift/Farbe, Host-Sitzung, Klavier). Der sofortige Wechsel zwischen Schreibmaschine und Zeigege-
rät ermöglicht eine schnellere Platzierung des Cursors.

Statt auf Schreibtischebene kann die virtuelle Tastatur bezogen auf die Positionen der Fingerspitzen definiert werden. Das heißt, daß die Taste für jeden Finger zugeordnet wird, bevor die Taste tatsächlich angeschlagen wird. Eine kurze, schnelle Abwärtsbewegung wählt dann die Eingabetaste aus. Die Tastenzuordnung vor dem Anschlag bietet einige einzigartige Merkmale für die Rückmeldung an den Bediener:

- Die Tastenzuordnung kann auf dem Bildschirm angezeigt werden (d.h. als Farbcodierung auf einem Tastaturanordnungs-Fenster).
- Elektronisch unterstützte und stabilisierte Tastenzuordnung

gen ermöglichen es, Bedienern zu helfen, die ihre Hände und Finger nur eingeschränkt bewegen können.

- Eine vergrößerte oder akustische Rückmeldung kann Bediener mit einer Sehbehinderung unterstützen.
- Die aktuelle Tastenfunktion (Umschaltstatus usw.) kann angegeben werden.
- Eine "abhebende" Fingerbewegung kann erkannt werden. Dies ist eine schnelle und äußerst intuitive Art, Tippfehler rückgängig zu machen.
- Die wirkliche Position der Fingerspitze bezogen auf die Tastenmitte kann überwacht werden (Farbe, Lautstärke des Anklickens und Abstand).
- Der Rechner kann die Mitte der virtuellen Tasten hinsichtlich der aktuellen Position der Finger und der Hand auf dynamische Weise anpassen. Ein fast stationärer Finger wird automatisch auf die virtuelle Taste "gelegt". Von dieser Stellung aus kann ein erfahrener Maschinenschreiber die anderen Tasten ohne optische Kontrolle finden.
- Die Bedienerführung ermöglicht es, eine Warnmeldung oder einen bestärkenden Hinweis zu geben, wenn besondere Tasten in Reichweite kommen.

Eine Tastenzuordnung vor dem Anschlag kann für eine effiziente Bedienerrückmeldung sorgen. Die Technik der virtuellen Tastatur kann verwendet werden, um Tastaturen in Standardgröße für Rechner zu erzeugen, die kleiner als ein Notebook sind. Andere Ausführungen reduzieren die Anforderungen an die Schreibtischoberfläche für PCs, geben ferne und geschützte Tastaturen an und lassen sich für Seh- und motorisch Behinderte anpassen.

Die Bild-(Handform-)Aufteilung wird sehr einfach, wenn ein Farb-

sensor verwendet werden kann und die Hände auf einer monochromen (blauen) Schreibtischschablone bewegt werden. Jedoch sollte eine Schwarz/Weiß-Abbildung genauso gut funktionieren, wenn sich die Arbeitsoberfläche durch eine Graustufe unterscheiden läßt, die sich wesentlich von der Tönung der menschlichen Haut abhebt. Wenn sich keine Vermutungen über den Schreibtischhintergrund anstellen lassen, ist eine Vorstufen-Bildverarbeitung erforderlich, um die Konturen der Hände des Bedieners festzustellen. Mit dem a-priori-Wissen, daß das Hintergrundbild konstant bleibt (Rauschen und Schattierung ausgenommen), und dem Wissen über die Hand als eine zusammenhängende Einheit, hält sich das notwendige Maß an zeitkritischer Verarbeitung in Grenzen.

Wie vorstehend erwähnt wurde, sollte eine Tastatur dem Benutzer Informationen über die Tasten und Rückmeldungen geben. Eine taktile Rückmeldung einer virtuellen Tastatur ist durch die Berührung der Fingerspitze mit der Arbeitsplatte gegeben. Eine akustische Rückmeldung (Klicks) kann von der Verarbeitungseinheit leicht erzeugt werden.

Informationen über die Tasten für eine virtuelle Tastatur können in einem Fenster auf der Monitoranzeige angezeigt werden. Schematische Darstellungen der Bilder der Hand können über einer Tastaturschablone angezeigt werden. Die gleichzeitige Anzeige der Tastatur mit der Anwendung ist aus ergonomischen Gründen vorteilhaft, da man den Blick zwischen Bildschirm und Tastatur nicht zu wechseln braucht (Hinweis: Dies ist dem heutigen Gebrauch eines Mauszeigers ganz ähnlich).

Das Erfassen der Fingerbewegungen kann dazu verwendet werden, Rückmelde-Informationen für den Bediener zu erzeugen, wenn der Tastenanschlag unnötig stark ist. Dies ermöglicht es, die Gefahr des durch starke Belastung bedingten RSI-Syndroms (tätigkeitsbezogene Verletzungen nach wiederholten Belastungen) zu reduzieren. Auf jeden Fall läßt sich die virtuelle Tastatur mit einer geringeren Belastung der Finger bedienen, da es keine physischen, gefederten Tasten gibt.

Eine virtuelle Tastatur ermöglicht eine individuelle Optimierung der Tastaturgröße. Dies ist absolut nicht nebensächlich. Die Größe der Hand eines Menschen variiert über einen weiten Bereich. Die Standard-Tastaturgröße ist ein Kompromiß. Wenn die Seh-, Tast- und Motorik-Funktionen des Bedieners zur Tastaturschnittstelle getrennt werden, können diese Funktionen zugunsten der Bedienerfreundlichkeit und folglich der Betriebsgeschwindigkeit optimiert werden.

Außerdem kann eine virtuelle Tastatur dupliziert oder einfach in getrennte Elemente für die rechte und die linke Hand aufgeteilt werden. Dies erlaubt eine äußerst entspannte Arbeitshaltung (z.B. in einem Stuhl mit Armlehne). Die Bedeutung für Behinderte ist offensichtlich.

Für Bediener mit weniger Sachkenntnis im Umgang mit Computern und weniger Erfahrung im Maschinenschreiben ist eine gedruckte Schablone (Platzvorlage) mit der Tastaturanordnung nützlich. Ein anderes Bildauswerteverfahren erleichtert es dann dem Bildverarbeitungssystem, die virtuelle Tastatur an die Schablone anzupassen. Die Platzvorlage kann auf ein Stück weiches Gewebe (z.B. Polyurethan-Schaum) aufgebracht werden, um die Belastung für die Fingerspitzen vom Anschlag auf einer harten Schreibtischplatte zu verringern.

A N S P R Ü C H E

1. Verfahren zur Erfassung von Benutzerbefehlen oder Daten in einem Befehls- oder Datenverarbeitungssystem, wobei vom Benutzer gemachte Bewegungen überwacht werden, das System einen TV-Sensor (2) enthält, der so plaziert ist, daß er einen unversperrten Blick auf die Hand und/oder die Finger ermöglicht, sowie Mittel zur Digitalisierung und Verarbeitung von Signalen von dem TV-Sensor,

wobei die Hand und/oder die Finger des Benutzers von dem TV-Sensor überwacht werden,

Form, Position und aktive Bewegungen der Hand und/oder der Finger von dem TV-Sensor gleichzeitig erkannt werden und

von den Mitteln für die Signalverarbeitung analysiert werden,

wobei die Bewegungen als entsprechende Eingabedaten oder Befehle interpretiert werden,

dadurch gekennzeichnet, daß die aktiven Bewegungen Tastenanschlagsbewegungen umfassen,

wobei die Tastenanschlagsbewegungen von aufeinanderfolgenden Bildfeldern des TV-Sensors erkannt werden.

2. Verfahren nach Anspruch 1, wobei die Analyse der Form und der Position die Lokalisierung der Fingerspitzen einschließt.
3. Verfahren nach Anspruch 1 oder 2, wobei die aktiven Bewegungen horizontale und vertikale Bewegungen und Fingerspitzendynamik einschließen.

4. Verfahren nach jedem der vorausgehenden Ansprüche, wobei das Verarbeitungssystem des weiteren einen Monitor (1) umfaßt und zumindest ein Teil des Anzeigebereichs des Monitors dazu dient, zumindest schematisch Hand- oder Fingerbewegungen in einem entsprechenden Bereich (4) vor dem TV-Sensor (2) visuell darzustellen.
5. Verfahren nach Anspruch 4, wobei der Teil des Anzeigebereichs automatisch aktiviert wird, indem eine Hand in dem Bereich (4) vor dem TV-Sensor (2) positioniert wird.
6. Verfahren nach Anspruch 4 oder 5, wobei schematische Darstellungen der überwachten Hand, der Finger oder Fingerspitzen auf dem Monitor angezeigt werden.
7. Verfahren nach Anspruch 6, wobei die schematischen Darstellungen entsprechend der aktuellen Position der Hand, der Finger oder Fingerspitzen vor dem TV-Sensor auf dem Monitor plaziert und bewegt werden.
8. Verfahren nach jedem der vorausgehenden Ansprüche, wobei die Analyse der aktiven Bewegungen eine Klassifizierung durch Vergleich der festgestellten Bewegungen mit vorgegebenen und gespeicherten Bewegungsmustern und die Unterscheidung von nicht übereinstimmenden Bewegungen einschließt, so daß nur bestimmte Bewegungsarten akzeptiert werden.
9. Verfahren nach Anspruch 8, wobei die akzeptierten bestimmten Bewegungsarten als Befehle interpretiert werden.
10. Verfahren nach jedem der vorausgehenden Ansprüche 1 bis 9, wobei das Erkennen der Tastenanschlagsbewegungen die Tastenzuordnung vor dem Anschlag einschließt, indem die Fingerspitzendynamik überwacht wird.
11. Verfahren nach jedem der Ansprüche 4 bis 10, wobei der Mo-

nitor (1) des weiteren Symbole anzeigt, die Datenwerte oder Befehle darstellen, die direkt ausgewählt werden können.

12. Verfahren nach Anspruch 11, wobei die Symbole so gewählt und angeordnet sind, daß sie eine Tastatur darstellen.
13. Verfahren nach jedem der vorausgehenden Ansprüche, wobei die Analyse der aktiven Bewegungen die Unterdrückung von weniger bedeutenden Bewegungen, die z.B. durch Verzerrungen wie beispielsweise Vibrationen etc. hervorgerufen werden, durch Filterung einschließt.
14. Daten- oder Befehlseingabegerät, wobei Hand- oder Fingerbewegungen eines Benutzers erkannt und analysiert werden, wie in jedem der vorausgehenden Ansprüche dargelegt ist.
15. Datenverarbeitungssystem, das mindestens ein Eingabegerät einschließt, wie in Anspruch 14 dargelegt ist.

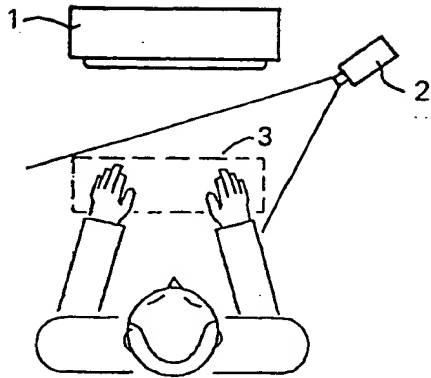


FIG. 1

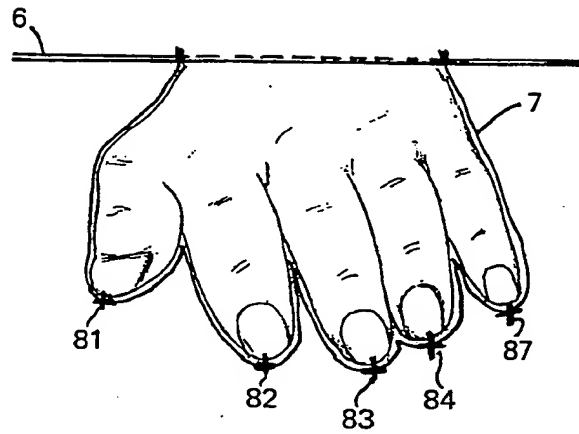


FIG. 3

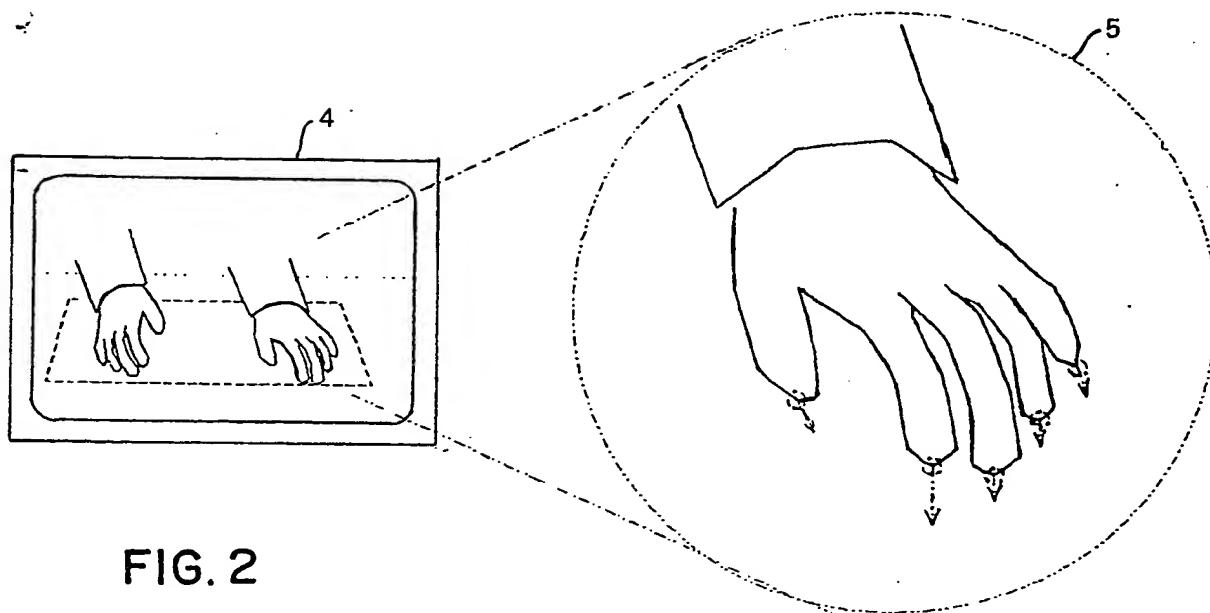


FIG. 2